WHAT IS CLAIMED IS:

1. A computer readable medium having a lexicon for storing word information and adapted for use by a language processing system, the lexicon comprising a plurality of entries, each entry corresponding to a word entered in the lexicon, wherein each entry comprises:

a first data field comprising spelling information for an entered word;

a second field comprising part of speech information associated with the entered word; and

a third field comprising lemma delta information associated with the entered word.

2. The computer readable medium of claim 1, the data fields further comprising:

a fourth field comprising descriptive information associated with the entered word; and

a fifth field comprising static segmentation mask information associated with the entered word.

3. The computer readable medium of claim 1, wherein the spelling information comprises an identification value corresponding to the entered word.

4. The computer readable medium of claim 1, wherein the first field further comprises dynamic segmentation information associated with the entered word.

5. The computer readable medium of claim 4, wherein the dynamic segmentation information comprises information for determining whether the entered word can be mapped to at least two separate lexical entries to recognize a valid compound term in a selected language.

6. The computer readable medium of claim 4, wherein the first field comprises up to 4 bytes of storage space.

7. The computer readable medium of claim 1, wherein the part of speech field comprises a part of speech for the entered word.

8. The computer readable medium of claim 7, wherein the part of speech field comprises a plurality of parts of speech associated with the entered word.

9. The computer readable medium of claim 8, wherein the part of speech field comprises up to four parts of speech, wherein each part of speech occupies up to 1 byte of storage space.

10. The computer readable medium of claim 8, and further comprising an intermediate indexes table

accessible by the language processing system, the intermediate indexes table comprising probability information for each of the parts of speech associated with the entered word.

11. The computer readable medium of claim 1, wherein the lemma delta information comprises transformation information associated with the entered word, the transformation information comprising information related to converting the entered word into a second word.

12. The computer readable medium of claim 11, wherein the second word is a lemma corresponding to the entered word.

13. The computer readable medium of claim 11, wherein the transformation information comprises an op code and an argument value.

14. The computer readable medium of claim 13, wherein the transformation information comprises up to four op codes and corresponding argument values.

15. The computer readable medium of claim 2, wherein the description information comprises up to 4 bytes of information.

16. The computer readable medium of claim 2, wherein the description information comprises named entity information.

17. The computer readable medium of claim 16, wherein the named entity comprises a proper pronoun.

18. The computer readable medium of claim 2, wherein the description information comprises information for at least one of person, tense, number, and gender associated with the entered word.

19. The computer readable medium of claim 2, wherein the static segmentation mask information comprises at least one constituent word length of the entered word, the entered word being a compound term of two or more constituent words.

20. A language processing system comprising the computer readable medium of claim 2.

21. The language processing system of claim 20, wherein the language processing system comprises an expansive stemming system.

22. The language processing system of claim 21, wherein the lexicon comprises data adapted for the expansive stemming system.

23. The language processing system of claim 20, wherein the lexicon comprises data adapted for a spell checker.

24. The language processing system of claim 20, wherein the lexicon comprises data adapted for a grammar checker.

25. The language processing system of claim 20, wherein the lexicon comprises data adapted for a speech recognition system

26. The language processing system of claim 20, wherein the lexicon comprises data adapted for a handwriting recognition system

27. The language processing system of claim 20, wherein the lexicon comprises data adapted for a machine translation system.

28. A lexicon stored on a computer readable medium, the lexicon comprising information for entered words, wherein for each entered word, corresponding word information is stored in data fields, the data fields comprising:

a spelling and dynamic segmentation field related to the entered word;

a part of speech field related to the entered word;

a lemma delta field related to the entered word;

a description field for the entered word; and

a static segmentation mask field for the entered word.

29. The lexicon of claim 28, wherein each field occupies up to 4 bytes of storage space.

30. A method of constructing a lexicon comprising information about words, for each word, the method comprising steps of:

storing spelling and dynamic segmentation information;

storing part of speech information; and

storing lemma delta information.

31. The computer readable medium of claim 30, and further comprising receiving lexical data comprising words to be entered into the lexicon.

32. The computer readable medium of claim 30, wherein receiving lexical data comprises receiving a dictionary of words.

33. The computer readable medium of claim 30, wherein receiving lexical data comprises receiving lexical data from at least one of web sources, newspapers, publications, and books.

34. The method of claim 31, and further comprising pre-processing the received lexical data to construct a word list of words to be entered into the lexicon.

35. The method of claim 30, and further comprising storing description information for each word.

36. The method of claim 35, and further comprising storing static segmentation mask information for words that are compound terms.

37. The method of claim 36, and further comprising storing information in a intermediate indexes table for some words in the lexicon.

38. The method of claim 37, wherein storing information comprises storing lemma delta information for the some words.

39. The method of claim 37, wherein storing information comprises storing probability information corresponding to parts of speech information for the some words.

40. A method of updating a lexicon, the method comprising the steps of claim 36, and further comprising:

selecting new words not currently in the lexicon and

storing information corresponding to the selected words to update the lexicon.